

Combinatorial approaches to carbohydrates

Daniel Kahne

Carbohydrates comprise an extremely important class of biological molecules but little is known about how they mediate their effects. This gap in understanding is largely due to the fact that obtaining pure carbohydrates in amounts large enough for biochemical studies is extremely difficult. The advent of combinatorial strategies to make carbohydrates promises to revolutionize the field of carbohydrate chemistry and biochemistry.

Addresses

Department of Chemistry, Princeton University, Princeton, NJ 08544, USA; e-mail: dkahne@princeton.edu

Current Opinion in Chemical Biology 1997, 1:130–135

<http://biomednet.com/elecref/1367593100100130>

© Current Biology Ltd ISSN 1367-5931

Introduction

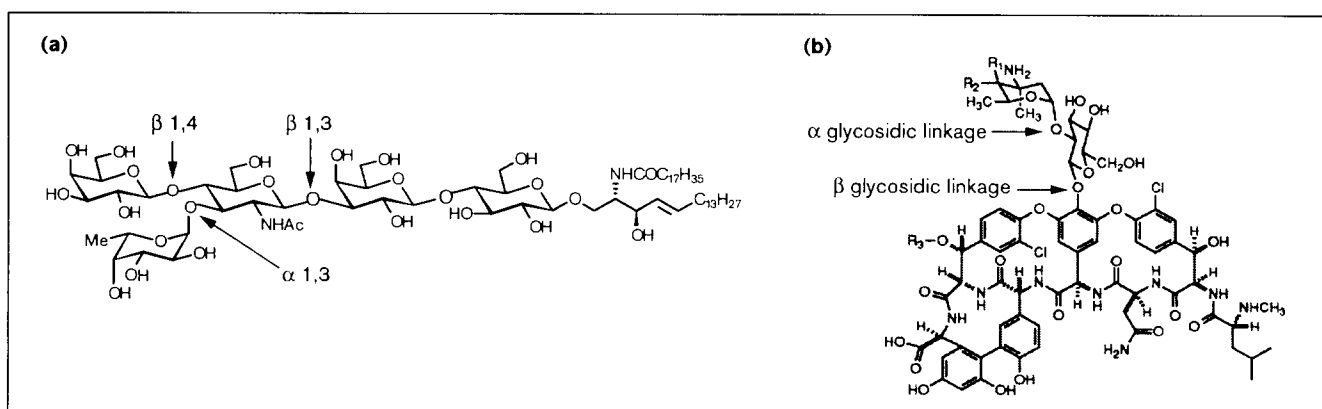
Carbohydrates are found widely in nature and perform a variety of important functions [1]. For example, oligosaccharides on the surfaces of eukaryotic cells mediate many fundamental cellular processes, including embryogenesis, tissue differentiation, inflammation, and metastasis. Cell surface carbohydrates also function as receptors for bacteria, viruses, and toxins [2]. Prokaryotic cells produce a variety of *O*-linked glycoconjugates with potent antitumor or antibiotic activity [3]. Some examples of natural carbohydrates are shown in Figure 1. A detailed understanding of the interactions between carbohydrates and their various ligands could lead to the ability to

influence important cellular recognition events, and to block or treat infection by microorganisms.

It is remarkable that carbohydrates have received so little attention given their importance. The single biggest reason for this apparent neglect is that carbohydrates are very difficult to synthesize efficiently. Biochemical studies on carbohydrate recognition require access to both the natural carbohydrate and a set of related analogues that can be used to probe the role of particular structural features in binding. Unfortunately, conventional strategies for carbohydrate synthesis are so time consuming that progress in studies on carbohydrate recognition has been painfully slow.

A combinatorial approach to the study of carbohydrates could dramatically accelerate progress in the field of carbohydrate recognition by making it possible to synthesize and screen hundreds to thousands of carbohydrates at a time. Although combinatorial chemistry has been used for several years to make peptide and nucleic acid libraries, researchers are only now beginning to apply combinatorial approaches to carbohydrates. Carbohydrates were first used as building blocks in glycopeptide and C-glycoside libraries; however, because glycosidic linkages were constructed individually prior to incorporation into the library [4–8], these libraries did not combinatorially explore the roles of the carbohydrates [9,10]. Combinatorial carbohydrate libraries have presented special difficulties because there are no general methods for the stereospecific chemical construction of glycosidic linkages. Moreover, because carbohydrates contain multiple hydroxyls, it is

Figure 1



Examples of carbohydrates. **(a)** Structure of Le^x , a natural cell surface oligosaccharide involved in both normal and pathogenic cell recognition processes. The shape of the molecule is largely determined by the glycosidic linkage conformation – anomeric stereochemistry and point of attachment. Modified from [32]. **(b)** Structure of vancomycin, a glycopeptide antibiotic. The attached oligosaccharide, which contains both α and β linkages, is critical for biological activity. Modified from [33].

necessary to have different protecting group schemes in order to make different types of glycosidic linkages. Only three approaches to the combinatorial synthesis of glycosidic linkages have been published to date, each presenting different solutions to the problems involved in the construction of carbohydrate libraries [11**,12–14,15**]. Approaches to the synthesis of carbohydrates and the future of combinatorial carbohydrate libraries are the subject of this review.

Solution combinatorial synthesis

Two of the approaches to the combinatorial construction of *O*-glycosidic linkages involve synthesis in solution. The first approach, developed by Hindsgaul and co-workers [11–13], involves coupling a protected glycosyl donor with a sugar acceptor containing several (3–5) free hydroxyls to produce a mixture of six to eight products (Fig. 2). Reactions were terminated at 30% completion to avoid *bis*-glycosylation of the acceptors. Remarkably, an approximately equal distribution of regioadducts was obtained in all the cases examined even though both primary and secondary hydroxyls were present in the glycosyl acceptors. The stereochemical outcome was dependent on the specific case, with 1:1 anomeric mixtures being obtained for some donors and α -linked products being obtained for others. After glycosylation the protecting groups were removed, and the coupled products were separated from starting materials using C18 chromatography.

Boons and co-workers [14] recently reported a somewhat more conventional approach to the combinatorial synthesis of glycosidic linkages in solution. To ensure the formation of regiospecific glycosidic linkages, they synthesized 10 protected disaccharide acceptors containing one free

hydroxyl each (Fig. 3). Construction of the building blocks for the library was highly convergent. The protected acceptors were combined and then reacted with a glycosyl donor under conditions that produced a 1:1 mixture of anomers. The resulting mixture contained an equal distribution of all twenty possible trisaccharides. The trisaccharide products were separated from the unreacted glycosyl donor by size exclusion chromatography.

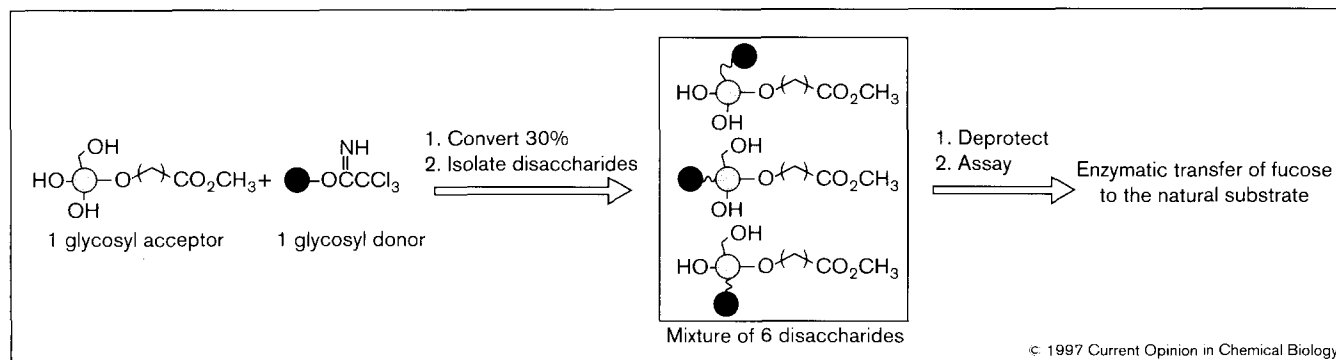
Solid-phase combinatorial synthesis

We recently reported a solid-phase approach to the combinatorial construction of glycosidic linkages [15**]. We constructed a carbohydrate library by coupling 12 different glycosyl donors to six different polymer-bound acceptors using a split and mix strategy (Fig. 4). The β linkages were constructed by using neighboring group participation; α linkages were constructed by carrying out the glycosylation reactions at low temperature in the absence of a participating group at C2. Thus, in a single combinatorial step, we made 72 different stereospecific glycosidic linkages. Additional diversity was introduced into the library by *N*-acylating the 72 products with an array of commercially available reagents to produce approximately 1300 different disaccharides and trisaccharides. To facilitate structure identification, the library was encoded at each combinatorial step using the method of Still and Wigler [16,17]. Readers interested in encoding strategies should refer to Czarnik's review on page 60 of this issue.

Screening and structure identification

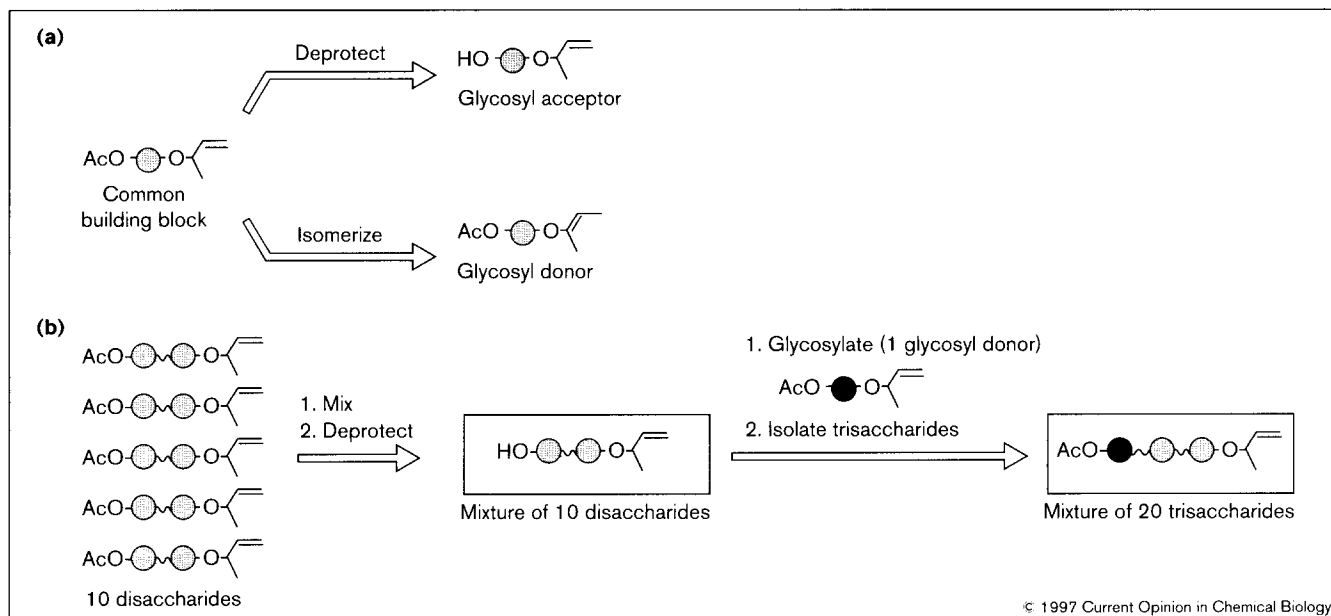
A combinatorial strategy for making glycosidic linkages is only the first step in the construction of a carbohydrate library. Since the point of making a library is to be able to identify rapidly a structure or set of structures with

Figure 2



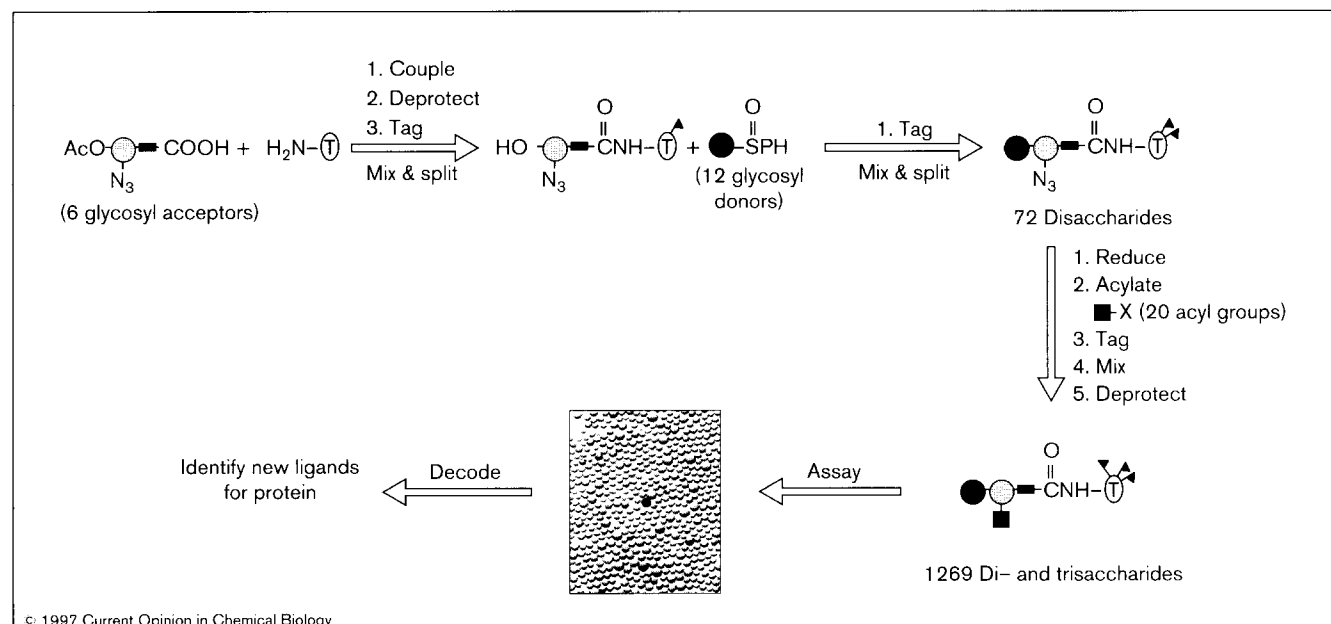
Hindsgaul's strategy for the rapid solution synthesis of carbohydrate libraries. A glycosyl acceptor containing three free hydroxyl groups was coupled with a protected glycosyl donor. Reactions were terminated at 30% completion to ensure that only one sugar monomer was added to each glycosyl acceptor. A fairly equal distribution of all the possible disaccharides was obtained in all of the cases examined. The stereochemical outcome was dependent on the specific case – shown is a 1:1 anomeric mixture of both α -linked and β -linked disaccharides (an equal distribution of six disaccharide products). After library synthesis, the disaccharides were deprotected, and separated from the starting materials using column chromatography. Hindsgaul's group went one step further, and screened this library for a substrate for a fucosyl transferase. Shaded circles represent different sugar monomers.

Figure 3



The Boon approach to the solution synthesis of carbohydrate libraries. (a) Both glycosyl acceptors and donors were synthesized from a common building block. (b) Library synthesis. Ten distinct protected disaccharides were mixed together and deacylated. These disaccharides contained one free hydroxyl. A glycosyl donor (a different sugar monomer) was added to the mixture and allowed to react. An equal distribution of α and β linkages was obtained, producing a library of 20 trisaccharides. Shaded circles represent different sugar monomers.

Figure 4



Kahne and coworkers' [15**] approach to the solid-phase synthesis of a combinatorial carbohydrate library. Using a split and mix strategy, 1269 disaccharides and trisaccharides were produced in three combinatorial steps. Each bead contained a single carbohydrate product and was encoded using the chemistry of Still and Wigler [16,17]. The beads were assayed in parallel for binding to a lectin using a colorimetric assay. The identities of the 'hit' carbohydrate ligands were determined by decoding the beads. Shaded circles represent different sugar monomers. The circled T represents a bead, squares represent different acyl groups, and the flags represent different (encoded) tags.

desirable activities, the utility of any synthetic approach for making a library cannot be fully evaluated without knowing how screening and structure identification will be carried out.

The reports on the solution approaches to the construction of glycosidic linkages have not addressed how screening will be carried out to identify novel structures. However, Hindsgaul and co-workers [12] showed that one of their reaction mixtures contained, as it was designed to, a substrate for a fucosyl transferase (Fig. 2). To identify the substrate, it was necessary to resynthesize all of the components of the mixture individually. Clearly, there is no advantage in synthesizing a mixture of compounds simultaneously if it is necessary to resynthesize a substantial fraction of the compounds in order to identify the active components; however, this work did demonstrate that it is possible to identify substrates in solution mixtures using enzymatic assays even when the substrate is a minor component of the mixture. It may be significantly more difficult to screen solution mixtures for binding activity, particularly if the mixtures contain an unequal distribution of products.

Our solid-phase library (mentioned above) was screened on-bead against biotin-labeled *Bauhinia purpurea* lectin, which is a model system for any carbohydrate-binding protein that recognizes surface-bound carbohydrates using polyvalent interactions. Binding of the lectin to individual beads was detected using a streptavidin-linked colorimetric enzyme assay [15••]. Remarkably, only a tiny fraction of the beads (0.3%) changed color in the assay. These beads were decoded, and two closely related carbohydrate ligands were identified as the best polyvalent binders for the lectin. This result validated the chemistry involved in making the library, since a failure in any of the synthetic steps would have produced many more beads containing similar structures. The fact that the lectin recognized two closely related carbohydrate ligands in a library containing 1300 carbohydrate ligands demonstrated, for the first time, that carbohydrate-binding proteins can discriminate exquisitely well between hundreds of different carbohydrates.

It is worth noting that the ligands identified in screening the solid-phase library contained a number of structural features which are unfavorable for binding when evaluated individually. These structural features apparently combine synergistically to produce a ligand that binds far better than expected. One implication of this finding is that deconvolution strategies may not have worked well for identifying the activity. Fixing one variable position at a time would not have worked because the avidity is not an additive function of the variables. Therefore, it would have been necessary to resynthesize a substantial fraction of the library to identify the activity. Carbohydrates are so difficult to synthesize that methods for structure identification which involve extensive resynthesis are not

appealing. Some direct methods of structure identification are also likely to be problematic because a large number of the individual components are positional isomers (diastereomers) and have the same mass. Therefore, encoding strategies may have particular utility for carbohydrate libraries. An alternative would be to synthesize the carbohydrate libraries in a parallel array format, where the identity of each compound is encoded by its location in the array.

Conclusions...

The construction of combinatorial carbohydrate libraries is still in its infancy. From a purely synthetic standpoint, the Hindsgaul approach to solution combinatorial synthesis is particularly appealing because unprotected glycosyl acceptors are used so the synthesis is fast. Approaches that use protected carbohydrate building blocks are time consuming because several steps are required to synthesize each building block. Library synthesis on the solid phase absolutely requires the use of protected monomers. Readers interested in synthesis on soluble polymers may refer to the review by Gravert and Janda on pp 107. Nevertheless, there are some important advantages to solid-phase carbohydrate synthesis [18–26,27••,28–30,31••] which makes it easier to utilize than a solution approach. One advantage is that reactions can be driven to completion by adding excess reagent; byproducts and excess starting materials can simply be washed away. Thus with solid-phase methods it is possible to produce single compounds on each bead with no need for purification. Screening and structure identification are greatly simplified when libraries are designed in such a way that the products are spatially resolved.

A special advantage of solid-phase carbohydrate libraries is that the components of the library can be screened either in solution after removal from the support or while still attached to the support. Because polymer-bound carbohydrates mimic the polyvalent presentation of cell surface carbohydrates, it is possible to screen the polymer-supported carbohydrate ligands for binding to proteins that utilize polyvalency in recognizing cell surface carbohydrates. Because these proteins bind weakly and non-selectively to monovalent carbohydrates in solution, it may be difficult, if not impossible, to identify good ligands for these proteins by screening solution carbohydrate libraries. Solid-phase presentation of carbohydrates could thus overcome this screening problem.

... and future directions

Two important issues relating to carbohydrate libraries have not been explicitly addressed in this review. One concerns the size of the carbohydrate structures that need to be made in order to identify interesting activity. This issue arises because many of the carbohydrate molecules found in nature are large, containing more than a dozen monosaccharide units. The synthesis of very large carbohydrate molecules is still an immense

undertaking. Fortunately, the recognition domains of most carbohydrate molecules range from one to about four sugars. Therefore, it should not be necessary to make libraries containing structures larger than tetrasaccharides. In fact, we have already demonstrated that even simple disaccharide libraries can contain ligands that bind with exquisite specificity to protein receptors [15**].

The second issue concerns the diversity of the carbohydrate libraries that can currently be made. The outcome (yield and stereoselectivity) of a glycosylation reaction is heavily dependent on the structures of both the glycosyl donor and acceptor. Clearly, it is impossible to make a library without a set of reliable bond-forming reactions which give predictable outcomes. Fortunately, there are now some glycosylation methods that give predictable outcomes for a wide range of donor-acceptor pairs. Therefore, although there are still some important unsolved problems in constructing glycosidic linkages, it is now possible to make diverse combinatorial carbohydrate libraries to study a wide range of recognition processes. In my opinion, the question now is no longer whether diverse carbohydrate libraries can be made, but how they should be used.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Varki A: **Biological roles of oligosaccharides: all of the theories are correct.** *Glycobiology* 1993, **3**:97-130.
2. Sharon N, Lis H: **Carbohydrates in cell recognition.** *Sci Am* 1993, **268**:82.
3. Nagarajan R: **Antibacterial activities and modes of action of vancomycin and related glycopeptides.** *Antimicrob Agents Chemother* 1991, 605-609.
4. Vetter D, Tumelty D, Singh SK, Gallop MA: **A versatile solid-phase synthesis of N-linked glycopeptides.** *Angew Chem Int Ed* 1995, **34**:60-63.
5. Peters S, Bielfeldt T, Meldal M, Bock K, Paulsen H: **Multiple-column solid-phase glycopeptide synthesis.** *J Chem Soc Perkin Trans* 1992, **1**:1163-1171.
6. Elofsson M, Roy S, Walse B, Kihlberg J: **Solid-phase synthesis and conformational studies of glycosylated derivatives of helper-T-cell immunogenic peptides from hen-egg lysozyme.** *Carbohydrate Res* 1993, **245**:89-103.
7. Roy R, Saha UK: **Rational design of multivalent glycoconjugate ligands. Synthesis of libraries of conformationally flexible rotamers of poly-N-linked lactosyl glycines.** *Chem Commun* 1996, 201-202.
8. Sutherlin DP, Stark TM, Hughes R, Armstrong RW: **Generation of C-glycoside peptide ligands for cell surface carbohydrate receptors using a four-component condensation on solid support.** *J Org Chem* 1996, **61**:8350-8354.
9. Patel TP, Goelz SE, Lobb RR, Parekh RB: **Isolation and characterization of natural protein-associated carbohydrate ligands for E-selectin.** *Biochemistry* 1994, **33**:14815-14824.
10. Mach H, Volkin DB, Burke CJ, Middaugh CR, Linhardt RJ, Fromm JR, Loganathan D: **Nature of the interaction of heparin with acidic fibroblast growth factor.** *Biochemistry* 1993, **32**:5480-5489.
11. Hindsgaul O: **A strategy of 'random glycosylation' for the production of oligosaccharide libraries.** *Angew Chem Int Ed* 1995, **34**:2720-2722.
•• An imaginative approach to carbohydrate libraries. Remarkably, glycosylation of unprotected acceptors containing both primary and secondary alcohols produces equivalent distribution of products.
12. Ding Y, Kanie O, Labbe J, Palcic MM, Ernst B, Hindsgaul O: **Synthesis and biological activity of oligosaccharide libraries.** In *Glycoimmunology*. Edited by Alavi A, Axford JS. New York: Plenum Press; 1995:261-269.
13. Ding Y, Labbe J, Kanie O, Hindsgaul O: **Towards oligosaccharide libraries: a study of the random galactosylation of unprotected N-acetylglucosamine.** *Bioorg Med Chem* 1996, **4**:683-692.
14. Boons GJ, Heskamp B, Hout F: **Vinyl glycosides in oligosaccharide synthesis: a strategy for the preparation of trisaccharide libraries based on latent-active glycosylation.** *Angew Chem Int Ed* 1996, **35**:2845-2847.
15. Liang R, Yan L, Loebach J, Ge M, Uozumi Y, Sekanina K, Horan N, Gildersleeve J, Thompson C, Smith A *et al.*: **Parallel synthesis and screening of a solid phase carbohydrate library.** *Science* 1996, **274**:1520-1522.
•• The first solid-phase combinatorial synthesis of a carbohydrate library. Parallel screening shows remarkable specificity for polyvalent carbohydrate-protein interactions.
16. Ohlmeyer MHJ, Swanson RN, Dillard LW, Reader JC, Asouline G, Kobayashi R, Wigler M, Still WC: **Complex synthetic chemical libraries indexed with molecular tags.** *Proc Natl Acad Sci USA* 1993, **90**:10922-10926.
17. Nestler HP, Bartlett PA, Still WC: **A general method for molecular tagging of encoded combinatorial chemistry libraries.** *J Org Chem* 1994, **59**:4723-4724.
18. Frecht JMJ: **Polymer-supported synthesis of oligosaccharides.** In *Polymer-Supported Reactions in Organic Synthesis*. Edited by Hodge P, Sherrington DC. New York: John Wiley & Sons; 1980:407.
19. Westerduin P, Veeneman GH, Pennings Y, van der Marel GA, van Boom JH: **Preparation of a fragment of the cell wall teichoic acid of *Bacillus licheniformis* ATCC 9945 via a solid phase approach.** *Tetrahedron Lett* 1987, **28**:1557-1560.
20. Veeneman GH, Notermans S, Liskamp RMJ, van der Marel GA, van Boom JH: **Solid-phase synthesis of a naturally occurring β -(1 \rightarrow 5)-linked D-galactofuranosyl heptamer containing the artificial linkage arm L-homoserine.** *Tetrahedron Lett* 1987, **28**:6695-6698.
21. Veeneman GH, Brugghe HF, van den Elst H, van Boom JH: **Solid-phase synthesis of a cell-wall component of *Haemophilus (Actinobacillus) pleuropneumoniae* serotype 2.** *Carbohydrate Res* 1990, **195**:C1-C4.
22. Verduyn R, van der Klein PAM, Douwes M, van der Marel GA, van Boom JH: **Polymer-supported solution synthesis of a heptaglycoside having phytoalexin elicitor activity.** *Recl Trav Chim Pays-Bas* 1993, **112**:464-466.
23. Danishefsky SJ, McClure KF, Randolph JT, Ruggeri RB: **A strategy for the solid-phase synthesis of oligosaccharides.** *Science* 1993, **260**:1307.
24. Randolph JT, McClure KF, Danishefsky SJ: **Major simplifications in oligosaccharide synthesis arising from a solid-phase method: an application to the synthesis of the Lewis b antigen.** *J Am Chem Soc* 1995, **117**:5712-5719.
25. Schuster M, Wang P, Paulson JC, Wong CH: **Solid-phase chemical-enzymatic synthesis of glycopeptides and oligosaccharides.** *J Am Chem Soc* 1994, **116**:1135-1136.

26. Yan L, Taylor CM, Goodnow R Jr, Kahne D: **Glycosylation on the Merrifield resin using anomeric sulfoxides.** *J Am Chem Soc* 1994, **116**:6953–6954.
27. Douglas SP, Whitfield DM, Krepinsky JJ: **Polymer-supported solution synthesis of oligosaccharides using a novel versatile linker for the synthesis of D-mannopentaose, a structural unit of D-mannans of pathogenic yeasts.** *J Am Chem Soc* 1995, **117**:2116–2117.
- Efficient and general strategy for synthesis of oligosaccharides on a soluble polymer. The approach combines advantages of both solid-phase and solution synthesis and may provide useful therapeutics.
28. Hunt JA, Roush WR: **Solid-phase synthesis of 6-deoxyoligosaccharides.** *J Am Chem Soc* 1996, **118**:9998–9999.
29. Ito Y, Kanie O, Ogawa T: **Orthogonal glycosylation strategy for rapid assembly of oligosaccharides on a polymer support.** *Angew Chem Int Ed* 1996, **35**:2510–2512.
30. Rademann J, Schmidt RR: **A new method for the solid phase synthesis of oligosaccharides.** *Tetrahedron Lett* 1996, **37**:3989–3990.
31. Nicolaou KC, Winssinger N, Pastor J, DeRoose F: **A general and highly efficient solid phase synthesis of oligosaccharides. Total synthesis of a heptasaccharide phytoalexin elicitor (HPE).** *J Am Chem Soc* 1997, **119**:449–450.
- A highly convergent synthesis of a large oligosaccharide on the solid phase. This work establishes that it is feasible – and possibly preferable in many cases – to make large oligosaccharides on a solid support.
32. Hakomori S, Zhang Y: **Glycosphingolipid antigens and cancer therapy.** *Chem Biol* 1997, **4**:97–104.
33. Solenberg PJ, Matsushima P, Stack DR, Wilkie SC, Thompson RC, Baltz RH: **Production of hybrid glycopeptide antibiotics *in vitro* and in *Streptomyces toyocaensis*.** *Chem Biol* 1997, **4**:195–202.